

2. Про електронні документи та електронний документообіг: Закон України № 851-IV від 22.05.2003 р. (редакція від 31.12.2023). Верховна Рада України. URL: <https://zakon.rada.gov.ua/rada/show/851-15> (Дата звернення: 01.05.2026)

3. PwC. Finance Effectiveness Benchmarking Report 2024 / AI agents help drive a new finance operating model. PricewaterhouseCoopers. URL: <https://www.pwc.com/us/en/tech-effect/ai-analytics/ai-agents-for-finance.html> (Дата звернення: 02.05.2026).

4. KPMG / Digital transformation statistics. WalkMe Blog, 2025. URL: <https://www.walkme.com/blog/digital-transformation-statistics/> (Дата звернення: 02.05.2026).

5. Юридична сила електронних документів в Україні. *Vchasno.ua*, 2025. URL: <https://vchasno.ua/en/yurydychna-syla-elektronnykh-dokumentiv/> (Дата звернення: 02.05.2026).

**Новіков М. В.,**

здобувач вищої освіти спеціальності 122 Комп'ютерні науки

**Науковий керівник:** Крайній В. О., к.е.н., старший викладач кафедри економічної кібернетики, комп'ютерних наук та інформаційних технологій, Миколаївський національний аграрний університет м. Миколаїв

## **МЕТОДИ АВТОМАТИЗОВАНОГО АНАЛІЗУ ТЕКСТІВ ДЛЯ ВИЯВЛЕННЯ МАНІПУЛЯЦІЙ У МЕДІА**

У тезах розглянуто сучасні методи автоматизованого аналізу текстів для виявлення маніпуляцій у медіа. Проаналізовано основні підходи обробки природної мови, включаючи токенізацію, лематизацію, методи векторного представлення тексту та машинне навчання для класифікації та аналізу новинних повідомлень. Особливу увагу приділено аналізу тональності тексту, семантичному аналізу та кластеризації новинних матеріалів для виявлення маніпулятивних наративів. Вказано на практичну значущість автоматизованих систем для медіамоніторингу, підвищення медіаграмотності та підтримки аналітичної діяльності у сфері медіакомунікацій.

У сучасному інформаційному суспільстві медіапростір відіграє ключову роль у формуванні громадської думки та впливі на соціально-політичні процеси. Швидке поширення інформації через новинні сайти, соціальні мережі та інші цифрові платформи створює умови для активного використання маніпулятивних технологій. Маніпуляції у медіа можуть проявлятися у вигляді перекручення фактів, емоційного впливу, вибіркового подання інформації або використання риторичних прийомів, спрямованих на формування певного ставлення аудиторії [1]. У зв'язку з великим обсягом медіаконтенту традиційні методи аналізу інформації стають недостатньо ефективними. Саме тому актуальним є застосування автоматизованих методів аналізу текстів, які дозволяють швидко

обробляти великі масиви даних і виявляти ознаки маніпуляцій у медіаповідомленнях [2].

**Мета дослідження** полягає у розгляді сучасних методів автоматизованого аналізу текстових даних та визначенні їх можливостей для виявлення маніпулятивних елементів у медіатекстах.

Автоматизований аналіз текстів ґрунтується на методах обробки природної мови (Natural Language Processing, NLP), які дозволяють комп'ютерним системам аналізувати, інтерпретувати та структурувати текстову інформацію. Основною метою таких методів є перетворення неструктурованого тексту на формалізовані дані, придатні для подальшого аналізу[3].

Першим етапом автоматизованого аналізу текстів є попередня обробка даних. Вона включає очищення тексту від зайвих символів, токенізацію, лематизацію та видалення стоп-слів. Токенізація передбачає поділ тексту на окремі слова або фрази, що дозволяє виділити базові елементи мовної структури. Лематизація забезпечує зведення слів до їхньої початкової форми, що сприяє підвищенню точності подальшого аналізу.

Важливим етапом є представлення текстових даних у вигляді числових векторів. Для цього використовуються різні методи, серед яких найбільш поширеними є модель «мішок слів» (Bag-of-Words), TF-IDF та сучасні методи векторного представлення слів[4]. Такі підходи дозволяють визначати значущість окремих слів у тексті та формувати математичні моделі для подальшої обробки.

Для виявлення маніпуляцій у медіатекстах активно застосовуються методи машинного навчання. Зокрема, використовуються алгоритми класифікації, які дозволяють автоматично розподіляти тексти за певними категоріями, наприклад «маніпулятивний» або «нейтральний». Серед найбільш поширених алгоритмів можна виділити наївний баєсівський класифікатор, метод опорних векторів, логістичну регресію та різні типи нейронних мереж. Ці алгоритми навчаються на основі попередньо розмічених текстів і можуть ефективно визначати характерні ознаки маніпулятивних повідомлень.

Одним із важливих напрямів автоматизованого аналізу є аналіз тональності тексту. Він дозволяє визначити емоційне забарвлення повідомлення та виявити використання надмірно емоційної або оцінної лексики, що часто є характерною ознакою маніпулятивних матеріалів [5]. Маніпуляції можуть також проявлятися через використання певних риторичних конструкцій, перебільшень або навмисного акцентування уваги на окремих аспектах подій.

Ще одним важливим інструментом є семантичний аналіз текстів. Він дозволяє досліджувати змістові зв'язки між словами та фразами, визначати тематику тексту та виявляти приховані смислові структури. Використання сучасних мовних моделей дає змогу більш точно аналізувати контекст повідомлень і визначати потенційні ознаки інформаційних маніпуляцій.

Крім того, у дослідженнях медіаконтенту використовуються методи кластеризації текстів. Вони дозволяють групувати новинні матеріали за подібними характеристиками, що сприяє виявленню інформаційних кампаній або повторюваних маніпулятивних наративів. Завдяки таким методам можна

ідентифікувати групи текстів, які мають схожі мовні особливості або подібні інформаційні акценти.

Автоматизований аналіз текстів також може поєднуватися з методами візуалізації даних. Графіки, тематичні карти та мережеві моделі дозволяють наочно представити результати аналізу та краще зрозуміти структуру інформаційних потоків у медіапросторі.

Застосування автоматизованих методів аналізу текстів має важливе практичне значення. Такі системи можуть використовуватися для моніторингу медіапростору, виявлення інформаційних маніпуляцій, аналізу новинних потоків та підтримки аналітичної діяльності у сфері медіадосліджень. Крім того, вони можуть сприяти підвищенню медіаграмотності та покращенню якості інформаційного середовища.

Отже, автоматизований аналіз текстів є ефективним інструментом для дослідження медіаконтенту та виявлення маніпулятивних повідомлень. Використання методів обробки природної мови, машинного навчання та семантичного аналізу дозволяє обробляти великі обсяги текстової інформації та визначати характерні ознаки маніпуляцій у медіа. Подальший розвиток таких технологій сприятиме створенню більш ефективних систем моніторингу інформаційного простору та підвищенню якості аналітичних досліджень у сфері медіакомунікацій.

### **Список використаних джерел**

1. Почепцов Г. Г. *Інформаційна політика та інформаційні війни* / Г. Г. Почепцов. Київ : Києво-Могилянська академія, 2016. 312 с.
2. Зернецька О. В. *Глобальний розвиток систем масової комунікації* / О. В. Зернецька. Київ : Наукова думка, 2018. 256 с.
3. Пелешишин А. М., Тимовчак-Максимець О. Ю. Методи аналізу текстової інформації в інформаційних системах / А. М. Пелешишин, О. Ю. Тимовчак-Максимець // Вісник Національного університету «Львівська політехніка». 2019. № 7. С. 45-56.
4. Манойло А. В., Трачук В. В. Інформаційно-аналітичні технології у дослідженні медіапростору / А. В. Манойло, В. В. Трачук // Наукові записки Інституту журналістики. 2020. Вип. 12. С. 78-91.
5. Додонов О. Г., Ландэ Д. В. Інформаційні технології аналізу текстових потоків / О. Г. Додонов, Д. В. Ландэ. Київ: Інститут проблем реєстрації інформації НАН України, 2017. 198 с.